Landscape of sialylation patterns identify biomarkers for diagnosis and prediction of response to anti-TNF therapy in crohn's disease

Chenglin Ye[1], Sizhe Zhu[2], Yuan Gao[1]* and Yabing Huang[1]*

[1]Department of Pathology, Renmin Hospital of Wuhan University, Wuhan, China, [2]Department of Otolaryngology-Head and Neck Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Crohn's disease (CD), a subtype of inflammatory bowel disease (IBD), causes chronic gastrointestinal tract inflammation. Thirty percent of patients do not respond to anti-tumor necrosis factor (TNF) therapy. Sialylation is involved in the pathogenesis of IBD. We aimed to identify potential biomarkers for diagnosing CD and predicting anti-TNF medication outcomes in CD. Three potential biomarkers (SERPINB2, TFPI2, and SLC9B2) were screened using bioinformatics analysis and machine learning based on sialylation-related genes. Moreover, the combined model of SERPINB2, TFPI2, and SLC9B2 showed excellent diagnostic value in both the training and validation cohorts. Importantly, a Sial-score was constructed based on the expression of SERPINB2, TFPI2, and SLC9B2. The Sial-low group showed a lower level of immune infiltration than the Sial-high group. Anti-TNF therapy was effective for 94.4% of patients in the Sial-low group but only 15.8% in the Sial-high group. The Sial-score had an outstanding ability to predict and distinguish between responders and non-responders. Our comprehensive analysis indicates that SERPINB2, TFPI2, and SLC9B2 play essential roles in pathogenesis and anti-TNF therapy resistance in CD. Furthermore, it may provide novel concepts for customizing treatment for individual patients with CD.

KEYWORDS

crohn's disease, anti-TNF therapy, sialylation, immune infiltration, bioinformatics analysis

Introduction

Crohn's disease (CD) is a significant type of inflammatory bowel disease (IBD) that is characterized by a chronic inflammatory condition that can affect any area of the gastrointestinal tract (Adegbola et al., 2018; Cushing and Higgins, 2021). Approximately 5% of people worldwide are affected by CD (Kaplan, 2015). In China, the estimated incidence rate is 0.51–1.09 per 100,000 people (Ng et al., 2017; Li et al.,

2019a). Currently, there are no curative treatments for CD. The tumor necrosis factor (TNF) inhibitor infliximab, the first biological response modifier, was licensed for treating CD in 1998 and increased patient response and remission rates (Adegbola et al., 2018). Up to 30% of patients do not respond to anti-TNF medications, and 50% of patients who initially benefit from these medicines lose clinical improvement within the first year, necessitating dosage increase or therapy change (Adegbola et al., 2018; Ye et al., 2022). Therefore, exploring effective therapeutic strategies for patients with CD is crucial.

Sialylation involves the addition of sialic acid to the terminal end of glycoproteins, a biologically significant alteration involved in microbial dysbiosis, gut inflammation, and immunological responses (Li and Ding, 2019; Giron et al., 2020). A recent study reported that sialylation of intestinal mucus by ST6GALNAC1 is essential for commensalism and bacterial metabolite balance as well as intestinal barrier integrity in IBD. The integrity of the mucus is preserved by ST6GALNAC1-mediated sialylation, which protected MUC2 from being degraded by certain bacterial-secreted mucinases (Yao et al., 2022). Meanwhile, a local release of free sialic acid during inflammation is probably facilitated by the increase in sialylation of intestinal mucins during colitis. This leads to an overgrowth of E. coli, which exacerbates the pro-inflammatory response by intestinal dendritic cells (Parker et al., 1995; Huang et al., 2015). However, a comprehensive analysis of multiple sialylation-related gene and their roles in CD is lacking. Therefore, exploring expression patterns and functions of sialylation-related gene may help to understand the heterogeneity and pathogenesis of CD.

This study comprehensively analyzed the expression patterns and functions of sialylation-related genes in CD using bioinformatics and machine learning. First, patients with CD were classified into two subtypes based on the expression of differentially expressed sialylation-related genes. The immune infiltration level and anti-TNF therapy response of patients with CD of the two subtypes were analyzed. Weighted gene co-expression network analysis (WGCNA), least absolute shrinkage and selection operator (LASSO) regression, random forest (RF), and support vector machine recursive feature elimination (SVM-RFE) were applied to further screen biomarkers for anti-TNF therapy response. Moreover, a scoring system, the Sial-score, was established to predict the response to anti-TNF therapy in patients with CD before or after treatment.

Method

Data collection

Using the search terms "Crohn's Disease and anti-TNF" or "infliximab", gene expression cohorts for CD were retrieved from the Gene Expression Omnibus database. The following cohorts were obtained: GSE16879 (213 CD inflamed tissues before and after infliximab treatment and 13 normal tissues) (Arijs et al., 2009), GSE111761 (lamina propria mononuclear cells isolated from six CD tissues) (Schmitt et al., 2019), GSE42296 (peripheral blood samples were obtained from 20 patients with CD) (Mesko et al., 2013), GSE107865 (whole blood samples were collected from 22 patients with CD) (Gaujoux et al., 2019), GSE102133 (55 inflamed CD tissues and 12 normal tissues) (Verstockt et al., 2019), and GSE179285 (47 inflamed CD tissues and 31 normal tissues) (Keir et al., 2021). The GSE16879, GSE111761, GSE42296, and GSE107865 cohorts contained clinical information on whether patients responded to infliximab treatment. According to the clinician's assessment, the patients were classified according to their response to infliximab based on endoscopic and histologic findings 6 or 14 weeks after the first infliximab treatment. Sialylation-related genes were obtained from the GeneCards database using the search term "sialyation". Genes with relevance scores >1 were selected for further analysis.
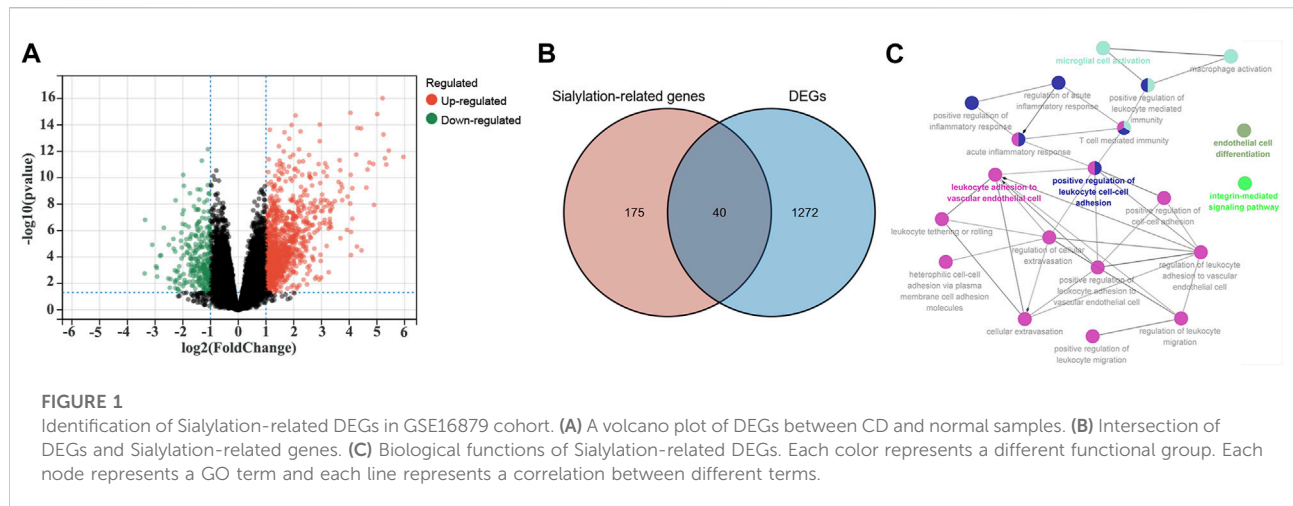
Differential analysis and unsupervised clustering

The Limma R package was used for differential analysis (Ritchie et al., 2015). Differentially expressed genes (DEGs) were classified as genes with an adjusted p-value <0.05 and | log2 (fold-change)| > 1.0. Sangerbox3.0 (http://vip.sangerbox.com/) was used to perform a consensus clustering algorithm using the R package ConsensusClusterPlus (Wilkerson and Hayes, 2010; Shen et al., 2022) to identify distinct subtypes. This was repeated 1,000 times to confirm clustering stability. Supplementary Table S1 shows the group information following the unsupervised clustering of cohorts.

WGCNA and single-sample gene set enrichment analysis

WGCNA was used to identify the related modules. The minimum number of module genes was set to 30, the parameter deepslip was set to 4, and the mergeCutHeight was set to 0.25. The hierarchical clustering dendrogram summarizes the gene modules with different colors (Langfelder and Horvath, 2008; Ye et al., 2021). Based on the cutoff criteria, genes with high connections in clinically relevant modules were identified as hub genes. The ClueGO plug-in was used to analyze biological functions in Cytoscape 3.8.2.

The relative level of immune cell infiltration was estimated using single-sample gene set enrichment analysis (ssGSEA). The gene signatures of the immune cells are listed in Supplementary Table S2 (Bindea et al., 2013).

FIGURE 1
Identification of Sialylation-related DEGs in GSE16879 cohort. **(A)** A volcano plot of DEGs between CD and normal samples. **(B)** Intersection of DEGs and Sialylation-related genes. **(C)** Biological functions of Sialylation-related DEGs. Each color represents a different functional group. Each node represents a GO term and each line represents a correlation between different terms.

Screening biomarkers based on machine learning

LASSO regression was applied to select potential biomarkers using the glmnet R package (Friedman et al., 2010). Binomial distribution variables were then used in the LASSO classification coupled with one standard error lambda value for the minimum criterion. RF, a tree-based ensemble of tree-structured classifiers, was created using the package "randomForest" using least error regression trees for clinical feature genes. The importance of the factors was ranked using "Mean Decrease Accuracy" and "Mean Decrease Gini". SVM is a type of generalized linear classifier that uses supervised learning to categorize binary data (Huang et al., 2018). SVM-RFE, a SVM-based algorithm, was applied to select relevant genes through nonlinear kernels (Sanz et al., 2018).

Construction of the predicting score system

The expression of potential biomarkers was used to develop a scoring system based on principal component analysis (PCA) to predict the response to anti-TNF medication. The Sial-score was calculated as follows: Sial-score = $\sum PC1i$, where i is the potential biomarker.

Results

Identification of sialyation-related DEGs

Differential expression analyses were performed using the limma R package in the GSE16879 cohort to identify sialylation-related DEGs between the CD and normal samples. A volcano plot of the DEGs is presented in Figure 1A. 215 sialylation-related genes were used to overlapped with DEGs, a total of 40 sialylation-related
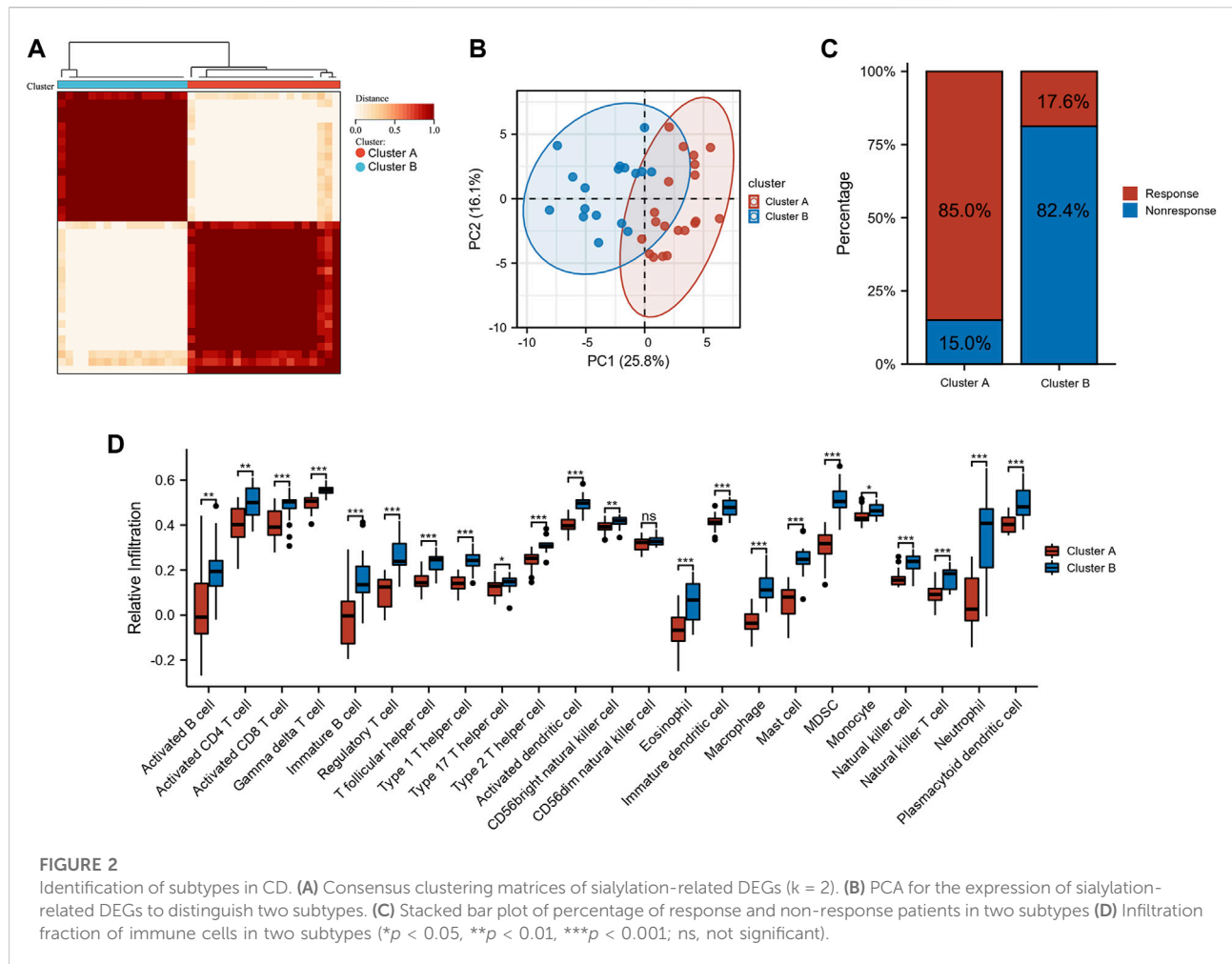
DEGs were obtained, including 38 upregulated and two downregulated (Figure 1B). Furthermore, ClueGO was used to explore the biological functions of the sialyation-related DEGs. As shown in Figure 1C, sialyation-related DEGs were significantly enriched in leukocyte adhesion to vascular endothelial cells, positive regulation of leukocyte cell-cell adhesion, microglial cell activation, endothelial cell differentiation, and integrin-mediated signaling pathways.

Unsupervised clustering for sialyation-related DEGs

Unsupervised clustering was used based on the expression files of the 40 sialylation-related DEGs. As shown in Figure 2A and Supplementary Figure S1, the two clusters had the best clustering effectiveness in the GSE16879 cohort. The transcription patterns of the sialyation-related DEGs between the two clusters differed significantly according to PCA (Figure 2B). Furthermore, we explored the number of patients who did or did not respond to anti-TNF therapy in clusters A and B. Patients in clusters A and B responded to anti-TNF therapy in proportions of 85.0% ($n = 20$) and 17.6% ($n = 16$), respectively (Figure 2C). ssGSEA was used to investigate the differences in immune infiltration between the two clusters. Figure 2D demonstrates significant differences between the two clusters of 28 immune cells; cluster B had a comparatively higher infiltration level than cluster A. Supplementary Table S3 lists the immune cell infiltration in the GSE16879 cohort.

Identification of potential biomarkers for predicting anti-TNF therapy response

Based on the excellent ability of subtypes to distinguish patients who had a response and nonresponse to infliximab,

FIGURE 2
Identification of subtypes in CD. **(A)** Consensus clustering matrices of sialylation-related DEGs (k = 2). **(B)** PCA for the expression of sialylation-related DEGs to distinguish two subtypes. **(C)** Stacked bar plot of percentage of response and non-response patients in two subtypes **(D)** Infiltration fraction of immune cells in two subtypes (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant).

we further screened the biomarkers for predicting anti-TNF therapy response. WGCNA was used to identify modules related to anti-TNF response. To build a scale-free network, the soft threshold β was set to 2 and no scale R^2 = 0.91. (Supplementary Figures S2A,B). Nine gene modules were identified, and the resulting gene dendrograms and module colors are presented in Supplementary Figure S2C. Figure 3A demonstrates that the blue module was negatively correlated with anti-TNF therapy response (module trait correlation = −0.54), whereas the green module had a positive correlation with anti-TNF therapy response (module trait correlation = 0.62). As patients in clusters A and B had different responses to anti-TNF treatment, the DEGs between clusters A and B were identified to explore the heterogeneity and characteristics of the two clusters (Figure 3B). Furthermore, to screen the genes related to anti-TNF treatment response, the genes in the blue and green modules were overlapped with DEGs between clusters A and B. As shown in Figure 3C, 35 and

19 overlapped genes were obtained, respectively. A total of 54 genes for further analyses. The biological functions of 54 intersecting genes were significantly enriched in metabolism-related pathways (Figure 3D).

54 intersecting genes were used to screen he potential biomarkers using machine learning. Twenty-three genes with clinical manifestations were selected using LASSO regression (Figures 4A,B). RF and SVM-RFE were used to identify genes related to the anti-TNF response. As shown in Figures 4C,D, seven genes were selected using both the RF and SVM-RFE algorithms. We analyzed the predictive abilities of 7 genes, SERPINB2 (area under the curve, AUC = 0.885), TFPI2 (AUC = 0.868), and SLC9B2 (AUC = 0.856), which could easily distinguish responders and non-responders before treatment (Figure 4E). As shown in Figures 4F–H, the expression levels of SERPINB2, TFPI2, and SLC9B2 are significantly downregulated in response group. Therefore, SERPINB2, TFPI2, and SLC9B2 were selected for further analyses.
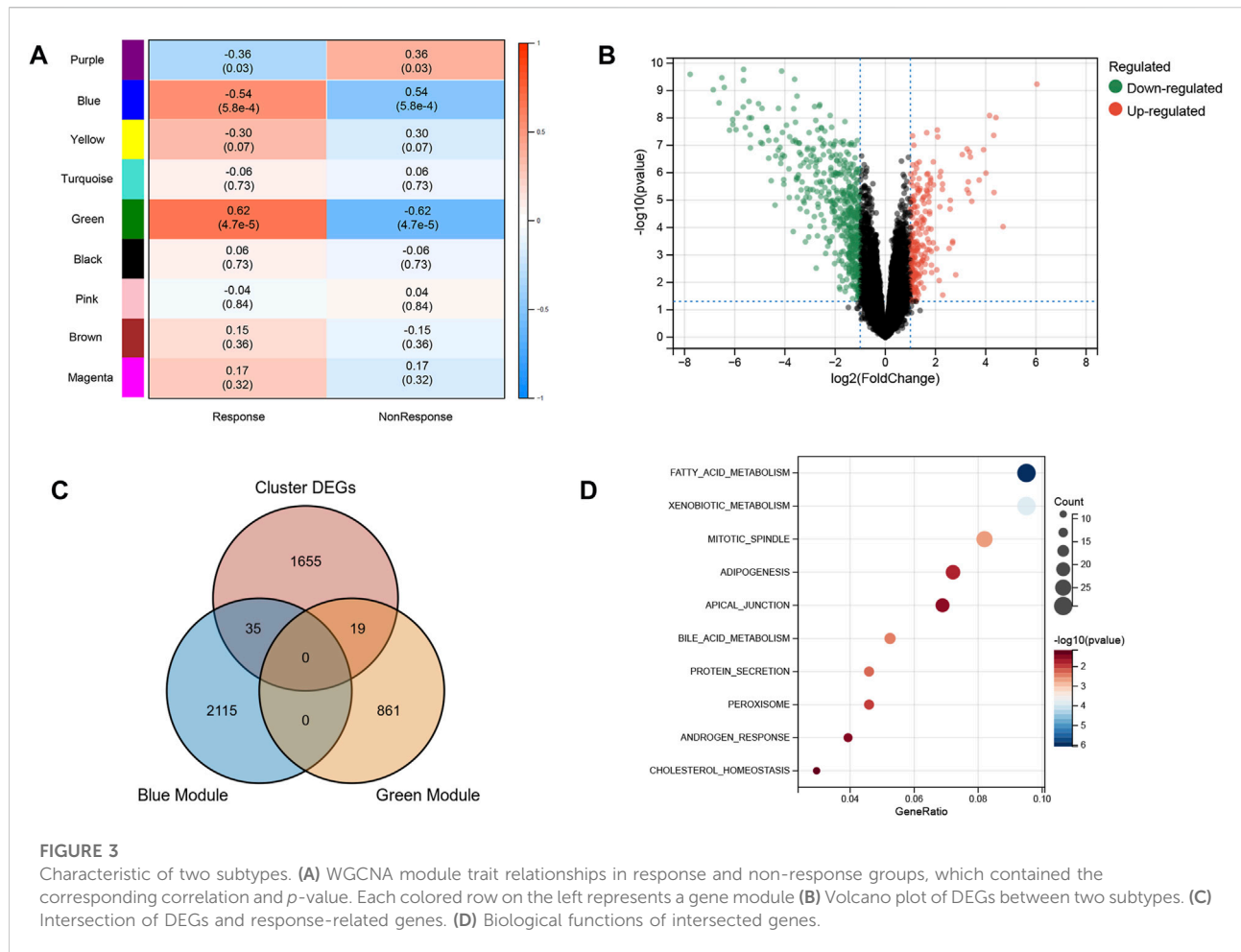
FIGURE 3
Characteristic of two subtypes. **(A)** WGCNA module trait relationships in response and non-response groups, which contained the corresponding correlation and *p*-value. Each colored row on the left represents a gene module **(B)** Volcano plot of DEGs between two subtypes. **(C)** Intersection of DEGs and response-related genes. **(D)** Biological functions of intersected genes.

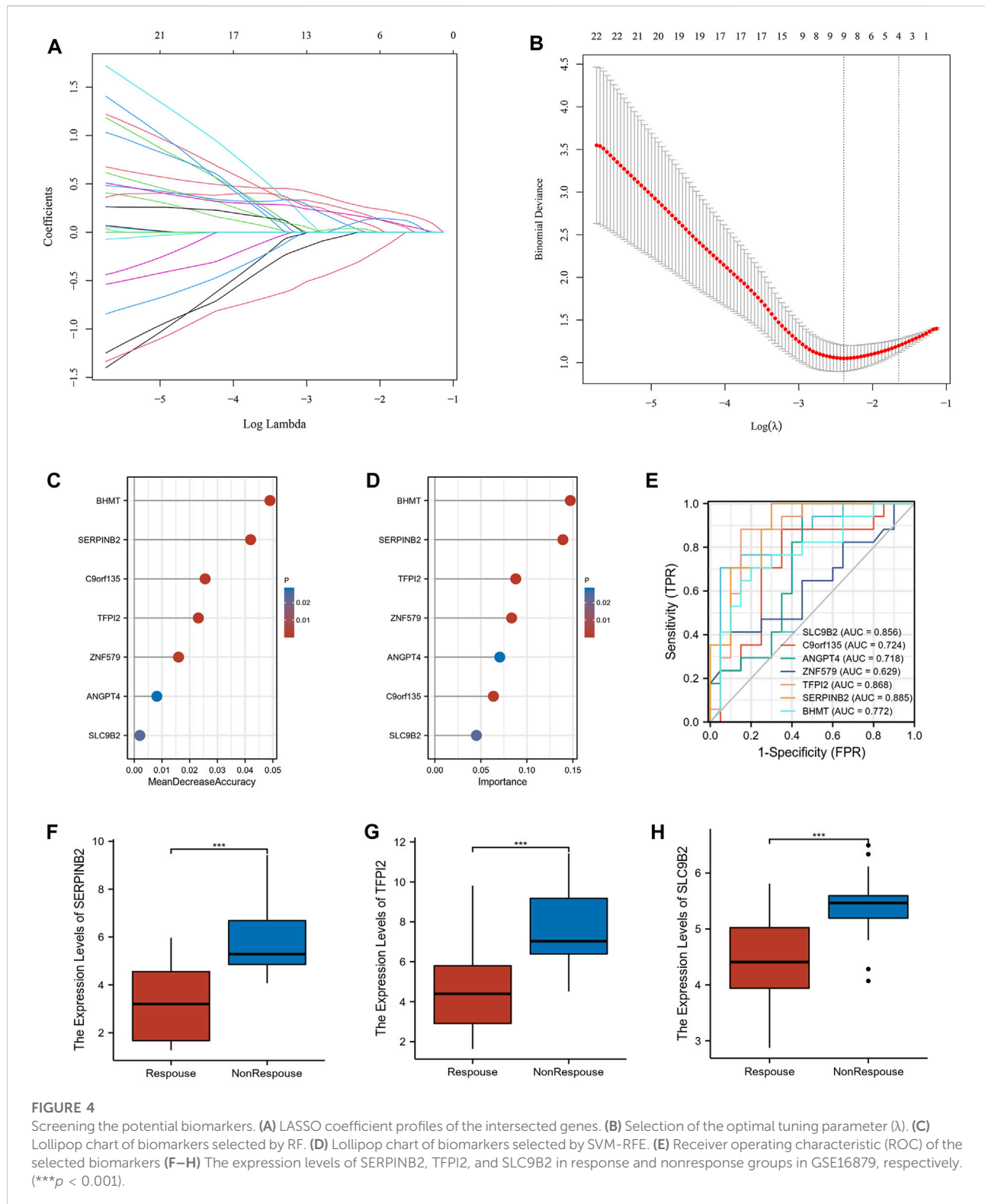## Diagnostic value of SERPINB2, TFPI2, and SLC9B2 in CD

The expression levels of SERPINB2, TFPI2, and SLC9B2 were significantly upregulated in CD (Figures 5A–C). We further constructed a combined model using logistic regression to explore the diagnostic value of SERPINB2, TFPI2, and SLC9B2. With an area under the curve (AUC) of 0.917, the ROC of the combined model demonstrated excellent discrimination for CD diagnosis (Figure 5D). In addition, the GSE179285 and GSE102133 cohorts were used to validate the diagnostic ability of the combined model. As shown in Figures 5E,F, the AUCs of the combined model in the GSE179285 and GSE94648 cohorts were 0.952 and 0.915, respectively.

Predictive value for Anti-TNF therapy response

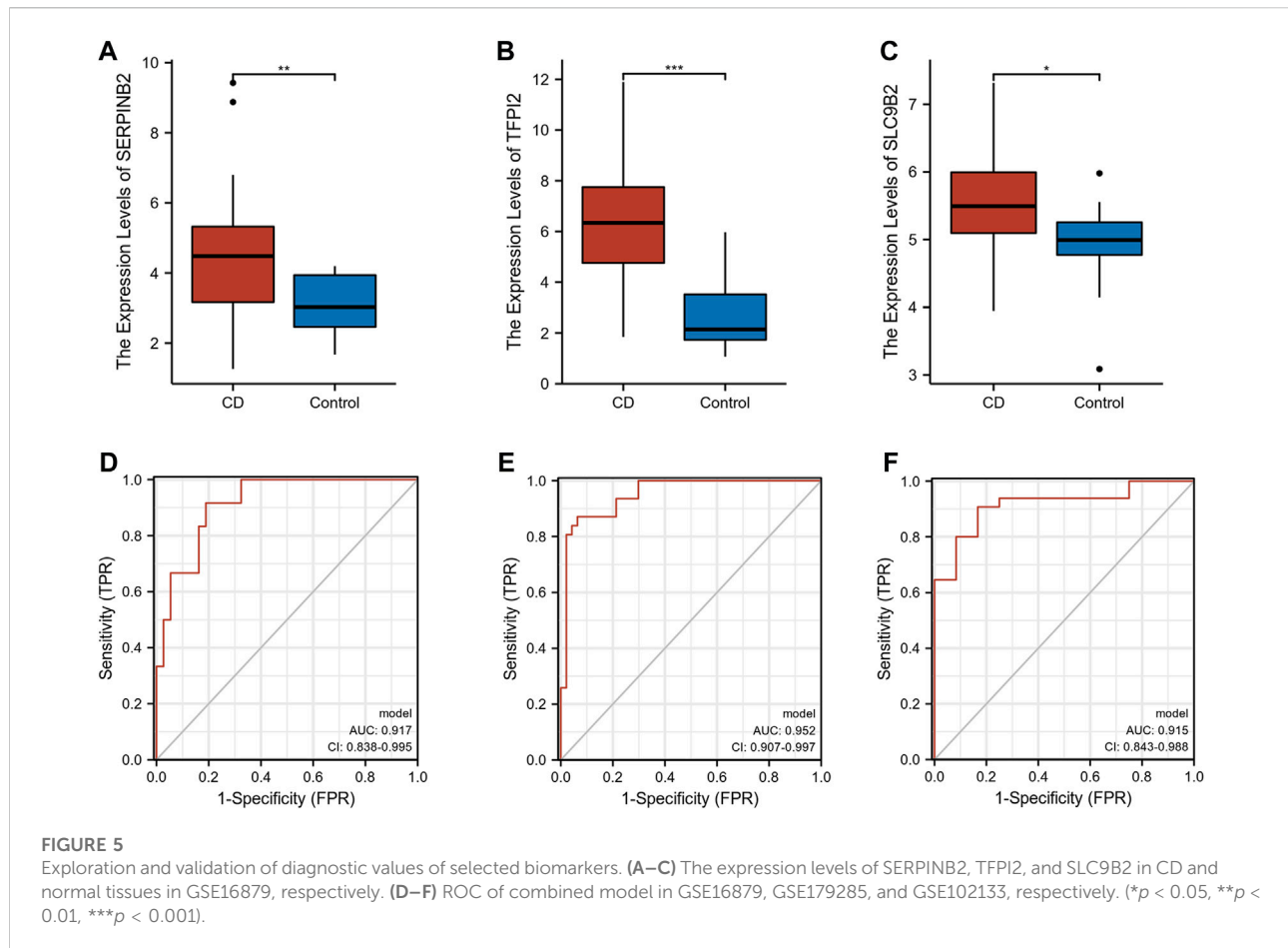Furthermore, a Sial-score based on the expression levels of SERPINB2, TFPI2, and SLC9B2 was constructed to predict anti-TNF responses. As shown in Figures 6A,B, the Sial-score had outstanding discrimination for responders and non-responders to anti-TNF therapy before (AUC = 0.912) or after (AUC = 0.920) treatment in the GSE16879 cohort. Subsequently, we collected three GEO cohorts to validate the predicted values of the Sial-score. As shown in Figure 6C, the Sial-score could easily distinguish responders from non-responders to anti-TNF therapy. Interestingly, the Sial-score could predict anti-TNF response in peripheral blood samples from patients with CD in the GSE422696 and GSE107865 cohorts (Figures 6D,E).

Patients with CD in GSE16879 were divided into Sial-high and Sial-low groups based on the Sial-score. To explore the characteristics of the Sial-high and Sial-low groups, we performed GSEA using the REACTOME pathway database. As shown in Figures 6F–I, the Sial-low group was mainly associated with cholesterol biosynthesis. In contrast, the Sial-high group showed conspicuous enrichment in immune-related pathways such as interleukin 10, interleukin 4, and interleukin 13 signalings. As shown in Figure 6J, patients in the Sial-high and the Sial-low groups responded to anti-TNF therapy in 15.8% (*n* =

FIGURE 4
Screening the potential biomarkers. **(A)** LASSO coefficient profiles of the intersected genes. **(B)** Selection of the optimal tuning parameter (λ). **(C)** Lollipop chart of biomarkers selected by RF. **(D)** Lollipop chart of biomarkers selected by SVM-RFE. **(E)** Receiver operating characteristic (ROC) of the selected biomarkers **(F−H)** The expression levels of SERPINB2, TFPI2, and SLC9B2 in response and nonresponse groups in GSE16879, respectively. (***$p < 0.001$).

19) and 94.4% ($n = 18$) of the patients, respectively. Further analysis of the two groups with immune cell infiltration showed that the Sial-high group was enriched in immune cell infiltration

(Figure 6K). Figure 6L illustrates the distribution of patients in the two subtypes, two Sial-score groups, and the response status to anti-TNF therapy.

FIGURE 5
Exploration and validation of diagnostic values of selected biomarkers. **(A–C)** The expression levels of SERPINB2, TFPI2, and SLC9B2 in CD and normal tissues in GSE16879, respectively. **(D–F)** ROC of combined model in GSE16879, GSE179285, and GSE102133, respectively. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).
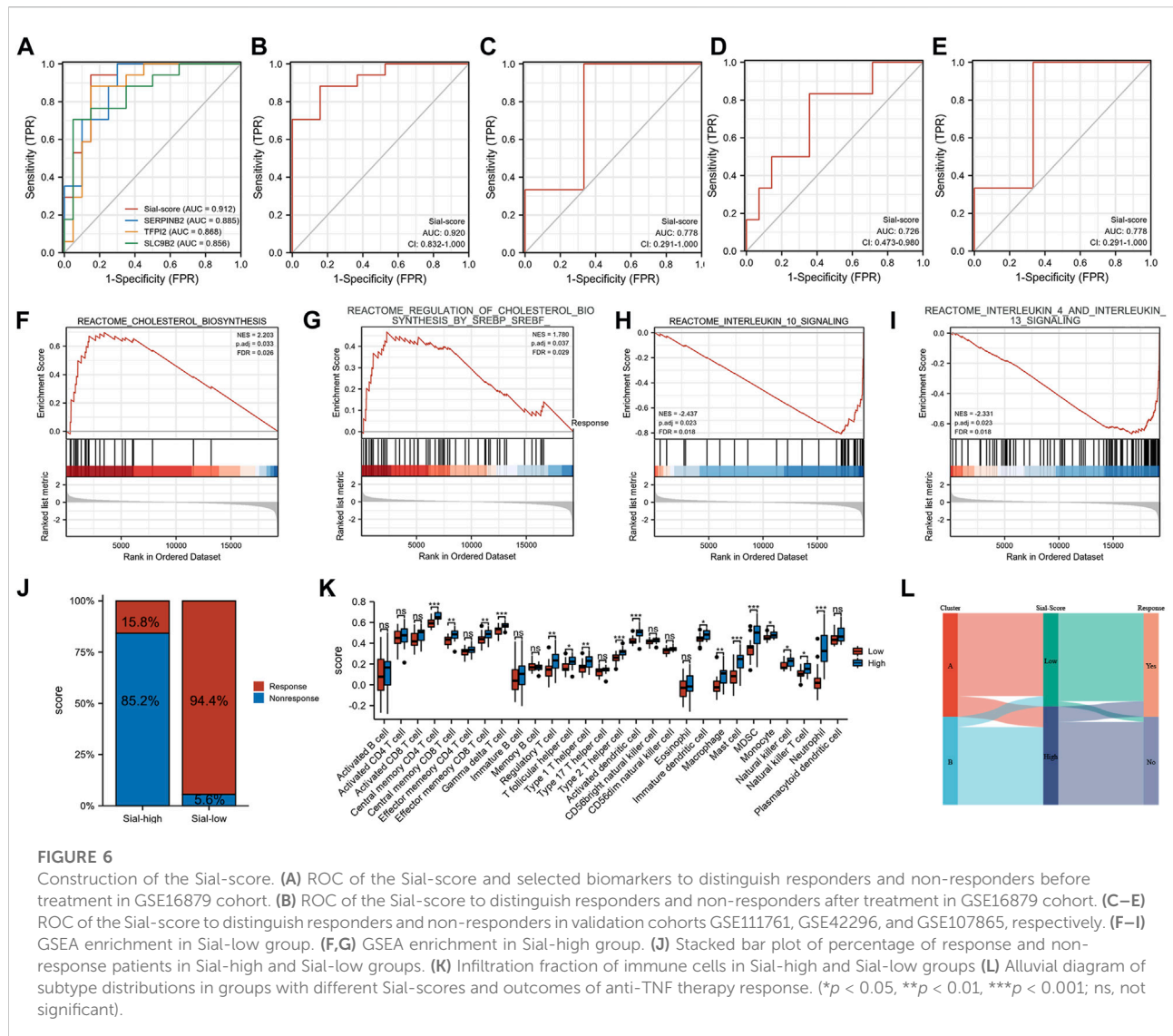
Discussion

Anti-TNF treatment helps patients with CD have better clinical outcomes, mucosal healing rates, and quality of life, although 10%–40% of individuals predominantly have no response (Cui et al., 2021). Pathophysiological heterogeneity has been a key factor limiting the outcome of new drug trials in patients with IBD over the past two decades (Bilsborough et al., 2016). However, identifying novel biomarkers is urgently needed to explore heterogeneity and provide personalized treatments for patients with CD.

Recently, Yao et al. (2022) reported that intestinal mucus sialylation by ST6GALNAC1 is critical for commensalism and bacterial metabolite homeostasis and that treatment with sialylated mucins reduces intestinal inflammation. Therefore, analyzing the expression patterns and functions of sialylation-related genes may benefit CD management. We collected data on sialylation-related genes and identified 40 genes that were differentially expressed in CD. To further explore the heterogeneity of CD, the two subtypes were analyzed using the unsupervised clustering method based on the expression of 40 sialylation-related genes. Anti-TNF therapy responders and

non-responders can be distinguished by two subtypes, with subtypes with more non-responders exhibiting higher levels of immune infiltration.

To screen biomarkers for predicting the outcome of anti-TNF medication, we identified two modules related to the anti-TNF response using WGCNA. Considering the unique characteristics of the two subtypes, DEGs between the two clusters were identified and overlapped with the two modules. Furthermore, SERPINB2, TFPI2, and SLC9B2 were selected as biomarkers of anti-TNF response using LASSO, RF, and SVM-RFE algorithms. Our results also indicated that the diagnostic model combining the expression of SERPINB2, TFPI2, and SLC9B2 showed excellent performance in both the training and validation cohorts.

SERPINB2, also known as plasminogen activator inhibitor 2, is highly expressed in peripheral blood from CD patients (Burczynski et al., 2006). Wei et al. (2015) reported upregulated SERPINB2 might serve as a target gene of downregulated miR-205 to activate inflammatory signal pathways in CD rat model. In several types of cancer, including colorectal cancer, TFPI2 has been identified as a tumor suppressor gene (Li et al., 2019b). TFPI2 promoter robustly hypermethylated in the patients with

FIGURE 6
Construction of the Sial-score. **(A)** ROC of the Sial-score and selected biomarkers to distinguish responders and non-responders before treatment in GSE16879 cohort. **(B)** ROC of the Sial-score to distinguish responders and non-responders after treatment in GSE16879 cohort. **(C–E)** ROC of the Sial-score to distinguish responders and non-responders in validation cohorts GSE111761, GSE42296, and GSE107865, respectively. **(F–I)** GSEA enrichment in Sial-low group. **(F,G)** GSEA enrichment in Sial-high group. **(J)** Stacked bar plot of percentage of response and non-response patients in Sial-high and Sial-low groups. **(K)** Infiltration fraction of immune cells in Sial-high and Sial-low groups **(L)** Alluvial diagram of subtype distributions in groups with different Sial-scores and outcomes of anti-TNF therapy response. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns, not significant).

CD (Kim et al., 2020). The methylation rates of TFPI2 elevated with progression of disease in inflamed colon tissue from patients with IBD, it seem to be a potential risk marker for colitis-associated cancer (Gerecke et al., 2015). SLC9B2 belongs to SLC9 family, mainly act as Na^+/H^+ exchangers and present in epithelial cells of the small intestine (Fuster et al., 2008; Cao et al., 2019). SLC9B2 deficiency may have an indirect effect on insulin secretion by interfering with clathrin-mediated endocytosis in β-cells (Fuster and Alexander, 2014).

To improve the clinical significance of our study, we constructed a Sial-score system based on SERPINB2, TFPI2, and SLC9B2 expression. The Sial-score has an outstanding ability to predict and classify the response to anti-TNF therapy in patients with CD before or after treatment. Importantly, only 15.8% of patients responded to anti-TNF therapy in the Sial-high group. However, 94.4% of patients with a low Sial-score

responded to anti-TNF therapy. Moreover, the Sial-high group was significantly enriched in immune-related pathways and showed a high level of immune cell infiltration. This indicated that immune cell infiltration plays a role in anti-TNF resistance. Martin et al. created a cell module named GIMATS, composed of IgG plasma cells, inflammatory mononuclear phagocytes, activated T cells, and stromal cells. Cell module scores differed between non-responsive and responsive patients with CD. Monocyte-derived macrophages dominated lesions enriched in the GIMATS module (Martin et al., 2019). The failure of patients with CD to respond to the anti-TNF medication has been attributed to innate transcriptional dysregulation of monocytes, resulting in increased activation of pro-inflammatory pathways (Gaiani et al., 2020). Our results indicate that SERPINB2, TFPI2, and SLC9B2 play important roles in pathogenesis and resistance to anti-TNF therapy in CD.

However, our study has some limitations. First, these datasets lacked data regarding important clinical variables such as disease activity and duration, previous bowel resection, and smoking. Second, we used retrospective data from public databases for our research. Future prospective studies are needed to confirm our findings.

Conclusion

In summary, a valid diagnostic model and scoring system for predicting anti-TNF therapy response was constructed based on the expression levels of SERPINB2, TFPI2, and SLC9B2. Our findings may aid auxiliary diagnoses and provide personalized treatment strategies for patients with CD.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

Conceptualization, CY and YH; data curation, CY, SZ, and YG; methodology, SZ; writing—original draft, CY; writing—review and editing, YH.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1065297/full#supplementary-material

References

Adegbola, S. O., Sahnan, K., Warusavitarne, J., Hart, A., and Tozer, P. (2018). Anti-TNF therapy in Crohn's disease. *Int. J. Mol. Sci.* 19 (8), E2244. doi:10.3390/ijms19082244

Arijs, I., De Hertogh, G., Lemaire, K., Quintens, R., Van Lommel, L., Van Steen, K., et al. (2009). Mucosal gene expression of antimicrobial peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* 4 (11), e7984. doi:10.1371/journal.pone.0007984

Bilsborough, J., Targan, S. R., and Snapper, S. B. (2016). Therapeutic targets in inflammatory bowel disease: Current and future. *Am. J. Gastroenterol. Suppl.* 3 (3), 27–37. doi:10.1038/ajgsup.2016.18

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39 (4), 782–795. doi:10.1016/j.immuni.2013.10.003

Burczynski, M. E., Peterson, R. L., Twine, N. C., Zuberek, K. A., Brodeur, B. J., Casciotti, L., et al. (2006). Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J. Mol. Diagn.* 8 (1), 51–61. doi:10.2353/jmoldx.2006.050079

Cao, L., Yuan, Z., Liu, M., and Stock, C. (2019). (Patho-)Physiology of Na(+)/H(+) exchangers (NHEs) in the digestive system. *Front. Physiol.* 10, 1566. doi:10.3389/fphys.2019.01566

Cui, G., Fan, Q., Li, Z., Goll, R., and Florholmen, J. (2021). Evaluation of anti-TNF therapeutic response in patients with inflammatory bowel disease: Current and novel biomarkers. *EBioMedicine* 66, 103329. doi:10.1016/j.ebiom.2021.103329

Cushing, K., and Higgins, P. D. R. (2021). Management of Crohn disease: A review. *JAMA* 325 (1), 69–80. doi:10.1001/jama.2020.18936

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Fuster, D. G., and Alexander, R. T. (2014). Traditional and emerging roles for the SLC9 Na+/H+ exchangers. *Pflugers Arch.* 466 (1), 61–76. doi:10.1007/s00424-013-1408-8

Fuster, D. G., Zhang, J., Shi, M., Bobulescu, I. A., Andersson, S., and Moe, O. W. (2008). Characterization of the sodium/hydrogen exchanger NHA2. *J. Am. Soc. Nephrol.* 19 (8), 1547–1556. doi:10.1681/ASN.2007111245

Gaiani, F., Rotoli, B. M., Ferrari, F., Barilli, A., Visigalli, R., Carra, M. C., et al. (2020). Monocytes from infliximab-resistant patients with Crohn's disease exhibit a disordered cytokine profile. *Sci. Rep.* 10 (1), 12238. doi:10.1038/s41598-020-68993-1

Gaujoux, R., Starosvetsky, E., Maimon, N., Vallania, F., Bar-Yoseph, H., Pressman, S., et al. (2019). Cell-centred meta-analysis reveals baseline predictors of anti-TNFα non-response in biopsy and blood of patients with IBD. *Gut* 68 (4), 604–614. doi:10.1136/gutjnl-2017-315494

Gerecke, C., Scholtka, B., Lowenstein, Y., Fait, I., Gottschalk, U., Rogoll, D., et al. (2015). Hypermethylation of ITGA4, TFPI2 and VIMENTIN promoters is increased in inflamed colon tissue: Putative risk markers for colitis-associated cancer. *J. Cancer Res. Clin. Oncol.* 141 (12), 2097–2107. doi:10.1007/s00432-015-1972-8

Giron, L. B., Tanes, C. E., Schleimann, M. H., Engen, P. A., Mattei, L. M., Anzurez, A., et al. (2020). Sialylation and fucosylation modulate inflammasome-activating eIF2 Signaling and microbial translocation during HIV infection. *Mucosal Immunol.* 13 (5), 753–766. doi:10.1038/s41385-020-0279-5

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., and Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer Genomics. *Cancer Genomics Proteomics* 15 (1), 41–51. doi:10.21873/cgp.20063

Huang, Y. L., Chassard, C., Hausmann, M., von Itzstein, M., and Hennet, T. (2015). Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice. *Nat. Commun.* 6, 8141. doi:10.1038/ncomms9141

Kaplan, G. G. (2015). The global burden of IBD: From 2015 to 2025. *Nat. Rev. Gastroenterol. Hepatol.* 12 (12), 720–727. doi:10.1038/nrgastro.2015.150

Keir, M. E., Fuh, F., Ichikawa, R., Acres, M., Hackney, J. A., Hulme, G., et al. (2021). Regulation and role of αE integrin and gut homing integrins in migration and retention of intestinal lymphocytes during inflammatory bowel disease. *J. Immunol.* 207 (9), 2245–2254. doi:10.4049/jimmunol. 2100220

Kim, T. O., Han, Y. K., and Yi, J. M. (2020). Hypermethylated promoters of tumor suppressor genes were identified in Crohn's disease patients. *Intest. Res.* 18 (3), 297–305. doi:10.5217/ir.2019.00105

Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Li, F., and Ding, J. (2019). Sialylation is involved in cell fate decision during development, reprogramming and cancer progression. *Protein Cell* 10 (8), 550–565. doi:10.1007/s13238-018-0597-5

Li, M., Tang, Y., Li, Q., Xiao, M., Yang, Y., and Wang, Y. (2019). Mono-ADP-ribosylation of H3R117 traps 5mC hydroxylase TET1 to impair demethylation of tumor suppressor gene TFPI2. *Oncogene* 38 (18), 3488–3503. doi:10.1038/s41388-018-0671-8

Li, Y., Chen, B., Gao, X., Hu, N., Huang, M., Ran, Z., et al. (2019). Current diagnosis and management of Crohn's disease in China: Results from a multicenter prospective disease registry. *BMC Gastroenterol.* 19 (1), 145. doi:10.1186/s12876-019-1057-2

Martin, J. C., Chang, C., Boschetti, G., Ungaro, R., Giri, M., Grout, J. A., et al. (2019). Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. *Cell* 178 (6), 1493–1508. doi:10.1016/j.cell.2019.08.008

Mesko, B., Poliska, S., Vancsa, A., Szekanecz, Z., Palatka, K., Hollo, Z., et al. (2013). Peripheral blood derived gene panels predict response to infliximab in rheumatoid arthritis and Crohn's disease. *Genome Med.* 5 (6), 59. doi:10.1186/gm463

Ng, S. C., Shi, H. Y., Hamidi, N., Underwood, F. E., Tang, W., Benchimol, E. I., et al. (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: A systematic review of population-based studies. *Lancet* 390 (10114), 2769–2778. doi:10.1016/S0140-6736(17)32448-0

Parker, N., Tsai, H. H., Ryder, S. D., Raouf, A. H., and Rhodes, J. M. (1995). Increased rate of sialylation of colonic mucin by cultured ulcerative colitis mucosal explants. *Digestion* 56 (1), 52–56. doi:10.1159/000201222

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Sanz, H., Valim, C., Vegas, E., Oller, J. M., and Reverter, F. (2018). SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinforma.* 19 (1), 432. doi:10.1186/s12859-018-2451-4

Schmitt, H., Billmeier, U., Dieterich, W., Rath, T., Sonnewald, S., Reid, S., et al. (2019). Expansion of IL-23 receptor bearing TNFR2+ T cells is associated with molecular resistance to anti-TNF therapy in Crohn's disease. *Gut* 68 (5), 814–828. doi:10.1136/gutjnl-2017-315671

Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1, e36. doi:10.1002/imt2.36

Verstockt, S., De Hertogh, G., Van der Goten, J., Verstockt, B., Vancamelbeke, M., Machiels, K., et al. (2019). Gene and mirna regulatory networks during different stages of Crohn's disease. *J. Crohns Colitis* 13 (7), 916–930. doi:10.1093/ecco-jcc/jjz007

Wei, K., Zhang, D., Hong, J., Zhang, C., Feng, X., Huang, Y., et al. (2015). Herb-partitioned moxibustion and the miRNAs related to Crohn's disease: A study based on rat models. *Evid. Based. Complement. Altern. Med.* 2015, 265238. doi:10.1155/2015/265238

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Yao, Y., Kim, G., Shafer, S., Chen, Z., Kubo, S., Ji, Y., et al. (2022). Mucus sialylation determines intestinal host-commensal homeostasis. *Cell* 185 (7), 1172–1188.e28. doi:10.1016/j.cell.2022.02.013

Ye, C., Zhu, S., and Yuan, J. (2022). Characterization of two TNF-related subtypes predicting infliximab therapy responses in Crohn's disease. *Front. Immunol.* 13, 871312. doi:10.3389/fimmu.2022.871312

Ye, C., Zhu, S., and Yuan, J. (2021). Construction of ceRNA network to reveal potential biomarkers in Crohn's disease and validation in a TNBS induced mice model. *J. Inflamm. Res.* 14, 6447–6459. doi:10.2147/JIR.S338053